

# 상태 오토마타와 기본 요소분류기를 이용한 가상현실용 실시간 인터페이스상

김 종 성<sup>†</sup> · 이 찬 수<sup>††</sup> · 송 경 준<sup>†</sup> · 민 병 의<sup>†</sup> · 박 치 항<sup>†††</sup>

## 요 약

본 논문에서는 가상현실의 기본 요소중의 하나인 사용자 인터페이스 분야에서 동적 손 제스처를 실시간으로 인식하는 시스템의 구현에 관하여 상술한다. 사람의 손과 손가락은 사람마다 같은 동작이라도 데이터의 변화가 다양하며 같은 동작을 반복해서 할 때에도 다른 데이터를 얻게되는등 시간에 따른 변화도 존재한다. 또한, 손가락의 외형 및 물리적 구조가 사람마다 다르기 때문에 다른 두사람에 의해 만들어진 같은 손 모양도 일반적인 센싱장비에 의해 측정될 때 다른 측정값을 나타낸다. 또한 동적 손제스처에서 동작의 시작과 끝을 명확히 구분하기가 매우 힘들다. 본 논문에서는 동적 손 제스처에 대해 각각의 의미있는 동작을 구분하기 위해 상태 오토마타를 이용하였고, 인식 범위의 확장성을 고려하여 동적 손 제스처를 퍼지 이론을 도입한 특징 해석에 의해 기본 요소인 손의 운동을 분류하고 퍼지 최대-최소 신경망을 적용하여 손의 모양을 분류함으로써 전체 손 제스처를 인식하는 시스템을 제안한다.

## Virtual Environment Interfacing based on State Automata and Elementary Classifiers

Jong-Sung Kim<sup>†</sup> · Chan-Su Lee<sup>††</sup> · Kyung-Joon Song<sup>†</sup>  
· Byungeui Min<sup>†</sup> · Chee-Hang Park<sup>†††</sup>

## ABSTRACT

This paper presents a system which recognizes dynamic hand gesture for virtual reality(VR). A dynamic hand gesture is a method of communication for human and computer who uses gestures, especially both hands and fingers. Since the human hands and fingers are not the same in physical dimension, the same form of a gesture produced by two persons with their hands may not have the same numerical values where obtained through electronic sensors. To recognize meaningful gesture from continuous gestures which have no token of beginning and end, this system segments current motion states using the state automata.

In this paper, we apply a fuzzy min-max neural network and feature analysis method using fuzzy logic for on-line pattern recognition.

---

† 정 회 원: 한국전자통신연구원 인공지능연구실  
 †† 정 회 원: 시스템공학연구소 가상현실연구실  
 ††† 정 회 원: 한국전자통신연구원 컴퓨터연구단  
 논문접수: 1997년 7월 30일, 심사완료: 1997년 10월 27일

## 1. 서 론

동적 손 제스처는 마우스나 키보드처럼 컴퓨터와 사람과의 상호 연결 수단중의 하나이다. 인간의 손은 주변의 환경과 상호 교류하기 위한 가장 유용하고 다양한 수단의 하나로써 특히, 종래에 키보드 또는 마우스를 이용하던 사람과 컴퓨터사이의 상호연결과정에 사람이 두손으로 표현하는 자유로운 몸짓을 직접 이용하려는 시도가 추후되고 있다. 손 운동 자체를 인식하는 분야의 연구는 크레인 조종을 위한 수신호 인식, 지휘자의 지휘동작 인식, 수화(sign language) 인식 등을 대상으로 미국, 일본 등지에서 활발히 추진되고 있다[1-5]. 특히, 가상현실(virtual reality) 연구 분야에서 사용하는 사용자 인터페이스에서 기존의 3차원 제어장치는 상당한 자유도를 가진 장비를 제어하기 위해 많은 전문성을 요구한다. 이와 같은 일들을 효과적으로 하기 위해 사용자와 기계가 3차원 정보를 교환하기 위한 새로운 입력장비가 필요하게 되었고 결국 사람의 손 제스처에 상당한 관심이 집중되었다[6].

사람이 손으로 표현하는 자세(gesture)는 손가락의 공간상의 고정된 모양으로 의미를 나타내는 정적인 자세(static posture)와, 손과 손가락의 공간적·시간적 움직임에 의해 의미를 나타내는 동적인 자세(dynamic gesture)로 구분된다.

본 논문에서는 사용자가 가상환경에서 특징의 손 제스처 명령을 수행하면 그 결과를 실시간으로 인식하는, 장갑장치에 기반을 둔 인식 시스템을 구현하고자 한다. 전체 시스템의 실시간성을 고려하여 본 논문에서는 손의 움직임을 감지하는 장비로 Virtual Technologies사의 CyberGlove™시스템[7]을 사용하였다.

정적인 손 자세를 인식할 때 사람마다 손의 크기 및 골격구조가 다르고 동적인 손의 운동의 경우에도 사람마다, 시간에 따라 매번 손의 운동 속도, 운동 크기 등이 다르므로 손의 자세를 전자적으로 감지한다는 측면에서 데이터 집합의 변화정도는 매우 심하다. 동적 손 제스처의 인식에 관한 연구는 미국 수화(American Sign Language)를 대상으로 몇 가지 연구가 있었다. Fels[3]는 미국수화에서 선택한 몇 가지 손 제스처를 인식하기 위해 역전파 신경망을 이용하였는데,

동작의 시작과 끝의 명확한 구분이 필요하며, 특히 많은 학습시간을 필요로 하는 시스템이다. 인공 신경망에서의 과도한 학습시간을 줄이기 위해 Waldron[5]은 자기구성 Map(Self Organizing Map)에 의한 재배열법을 제안하였다. 이 방법은 초기 집단의 인식을 위해 약간의 학습시간은 필요하지만 새로운 집단이 추가되었을 때의 과도한 학습시간을 줄일 수 있는 장점이 있지만 전체 인식율의 저하가 발생한다. 또한, 최근에 한글 수화(Korean Sign Language)에 대하여 학습 시간을 줄이면서 인식 단어의 확장이 쉬운 시스템을 제안한 바 있으나[16], 역시 연속적인 동적 동작에 대한 인식은 불가능 하다.

본 논문에서는 이와 같은 어려움을 극복하기 위하여 동적인 손의 움직임에서 의미 있는 제스처를 분리하기 위해 상태 오토마타를 설계하였고, 퍼지 이론을 이용한 손의 운동 방향 분류기와 P. Simpson[8]이 제안한 퍼지 최대-최소 신경망을 이용하여 손 모양 분류기를 개발하였다. 이 패턴 분류기는 새로운 집단(class)을 학습할 수 있고 기존 집단의 정보를 없애지 않으면서 온라인으로 빠르게 집단을 분류할 수 있었다.

2장에서는 본 논문의 인식 대상인 가상 현실에서의 가상교육용 손 제스처 명령들에 대한 특징, 해석 및 인식 알고리즘을 설명하고, 3장에서 인식 알고리즘을 포함한 손 제스처 인식 시스템의 구성을 설명한다. 4장에서는 실제 실험을 통하여 전체 시스템의 운영 결과를 보이며, 마지막으로 5장에서 결론을 맺는다.

## 2. 가상현실에서의 손 제스처 특징 해석

본 장에서는 인간의 동적 손 제스처를 실시간으로 인식하는 가상 현실용 인터페이스 기법에 관하여 상술한다. 가상현실 분야에서 인간과 컴퓨터의 상호 작용을 위한 도구로서 마우스, 키보드, 펜, 음성, 영상등 다양한 방법이 제안되고 있다. 본 논문에서는 인간이 가장 빨리 의사를 전달할 수 있고 다양한 표현이 가능한 손 제스처를 이용하였다. 즉, 가상공간상에 존재하는 가상의 물체(Avatar)를 움직이거나 주행(navigation)하는 일 등을 손 제스처 명령을 설정하여 실시간으로 제어하도록 한다. 가장 사용자에게 친숙한 인터페이스는 모든 사람이 이해하는 비언어적인(non-verbal) 통신 수단을 도입하는 것이지만 3차원 공간에

서 움직이는 영상의 고속 처리는 설령 비디오 카메라를 사용하더라도 매우 어렵다. 본 논문에서는 손 제스처의 데이터를 감지하는 장치로 최근에 가상현실 분야에서 주요한 입력 장치로 사용되고 있는 Virtual Technology사의 CyberGlove™[7] 시스템을 이용하였는데, 이는 각 손가락 관절의 굽힘 정도를 실시간으로 측정할 수 있다. 또한 장갑의 손등에는 손의 3차원 공간상의 자세(x, y, z, yaw, pitch, roll)를 측정하는 Polhemus센서가 부착되어 있다. 모두 13개의 손에 대한 정보(손가락 굽힘, 손의 위치)가 1/15초마다 측정되어지고 실시간 그래픽과 함께 인식된 결과가 컴퓨터에 표시된다. 본 논문의 좌표축은 CyberGlove™ 시스템과 관련된 것으로 수평방향의 좌, 우 운동은 x축, 수직방향의 상, 하 운동은 y축, 그리고 z축은 앞, 뒤 방향의 운동으로 정의되어 있다.

일반적으로 손의 움직임에 의해 표현되는 동작은 매우 다양하지만 이러한 많은 표현 가능한 동작을 체계적으로 정리한 것이 한글 수화(Korean Sign Language)라고 할 수 있다. 본 논문에서는 가상 현실용 손 제스처로 정의된 10가지 명령에 대한 동작들을 정의하기 위해 한글 표준 수화사전[9]을 참고하였다. (그림 1)은 본 논문에서 인식하고자 하는 손 제스처에 대한 10가지 명령을 보인 것이다. 여기서 ‘멈춤’은 가상 공간상의 제어 대상 물체(Avatar)를 정지하게 하는 것이고, ‘추적’은 현재의 손의 위치를 가상 공간상에서 똑같이 움직이도록 하는 것을 의미하며, ‘잡기’는 가상 공간상의 물체를 잡는 명령이다. 이상의 3가지 동작은 사용자가 임의의 방향으로 손을 움직일 수 있도록 하였다. 본 논문에서는 사용자가 명령을 내릴 때의 편의를 위해 손의 회전은 고려하지 않고 있으며

따라서 손 제스처는 다음과 같이 표현된다.

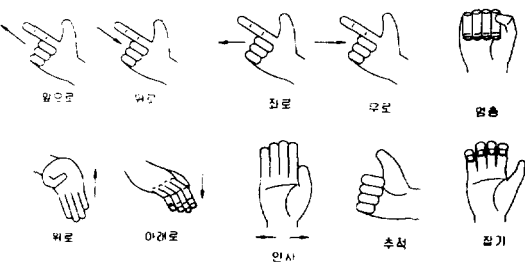
$$G_k = (x_k, y_k, z_k, t_k, F_{ki}), 1 \leq k \leq K: \text{time step}[1/5 \text{ sec.}]$$

$$F_{ki}: \text{flex angles of right hand}, i = 1, \dots, 10.$$

본 논문에서 다루고자하는 (그림 1)의 동적 손 제스처들에 대하여 실시간으로 인식하는 경우에 가장 먼저 고려하여야 하는 부분이 연속적인(continuous) 동적(dynamic) 동작에 대한 구분(segmentation)이다. 즉, 실시간으로 측정되는 손의 움직임에 대한 정보로부터 의미 있는 손 제스처와 의미 없는 손 제스처의 움직임을 구분할 수 있어야 한다. 본 논문에서는 연속된 제스처에서 개별 제스처를 분리하고 의미 없는 제스처를 제거하기 위하여 속력과 속력 변화에 따라 5가지 동작 위상을 구분하였으며, 이 구분된 동작 위상을 입력 사건으로 한 상태 오토마타에 의하여 5가지 동작 상태를 구분하였다.

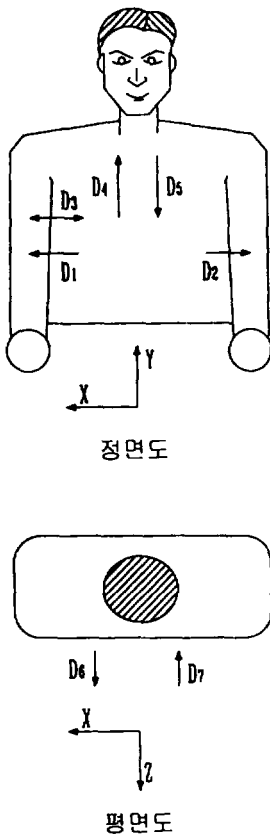
서론에서 설명한 기존의 동적 손 제스처 인식 시스템은 손 제스처 전체의 특징들을 입력으로 사용한 단일의 분류기를 사용하였지만, 본 연구에서는 동적 손 제스처를 두개의 기본적인 요소(fundamental components)로 분류하였는데, 즉, 손의 운동과 손의 모양에 의해 손 제스처를 구성할 수 있다. 본 논문에서는 복잡한 손 제스처를 손 운동 분류기(classifier)와 손 모양 분류기로 구분하여 두 개의 인식 결과를 결합하여 동적 손 제스처 인식기를 구성하고자 한다. 이와 같은 방법으로 전체 인식 시스템을 구성할 경우 새로운 손 제스처를 인식기에 추가할 경우 인식 동작의 확장이 쉽게 가능하다. (그림 1)을 인식기 측면에서 분석한 후 기본적인 손 운동 방향으로 (그림 2)와 같은 7가지의 방향 집단을 얻었고 손의 모양은 (그림 3)과 같이 모두 6종류의 집단이 존재한다. 앞에서 설명한 바대로 손의 모양은 본 시스템의 특성상 회전에 관계없이 같은 집단으로 분류하게 된다.

손 운동은 3차원 공간에서 이루어지기 때문에 사람마다 같은 동작이라도 변화가 다양하며 같은 동작을 반복해서 할 때에도 다른 데이터를 얻게되는 등 시간에 따른 변화도 함께 존재한다. 또한, 손가락의 외형 및 물리적 구조가 사람마다 다르기 때문에 다른 두사람에 의해 만들어진 같은 손 모양도 일반적인 센싱장비에 의해 측정될 때 다른 측정값을 나타낸다. (그림

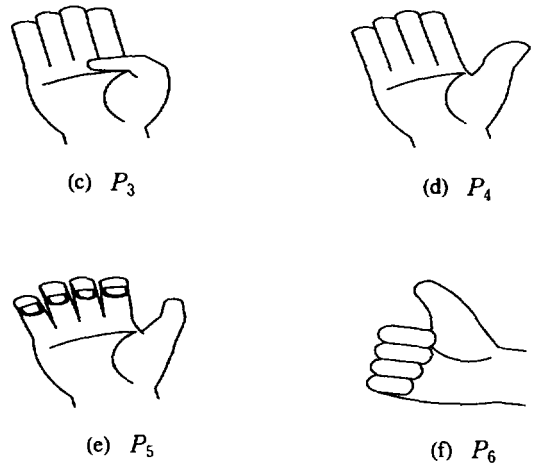
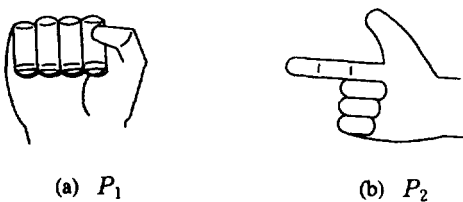


(그림 1) 10가지 손 제스처 명령 집단  
(Fig. 1) 10 gestures for VR

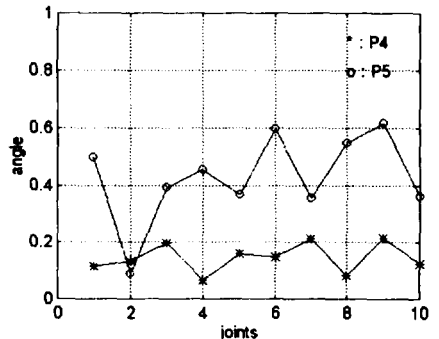
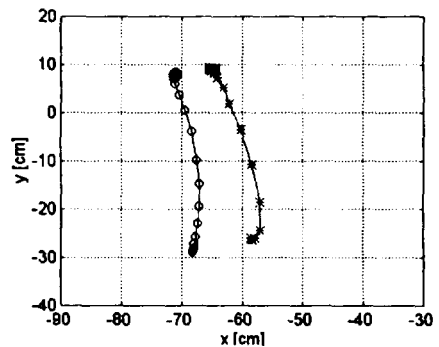
4(a)는 두사람에의해 센싱된 상향 운동( $D_4$ )의 궤적을 보인 것인데, 각각의 크기나 운동의 속도('\*'와 'o'로 표시된 것은 샘플된 데이터)가 다름을 알 수 있다. (그림 4(b))는 각기 다른 두 종류의 손의 모양( $P_4, P_5$ )에 대한 센싱 데이터를 나타낸 것인데 각 손가락마다 2개씩의 관절(첫번째, 두번째)에 대한 굽힘 정보를 0과 1사이로 정규화(normalization)하여 표시한 것이다.



(그림 2) 7가지의 기본 방향 집단  
(Fig. 2) 7 Basic direction classes



(그림 3) 6가지의 기본 손 모양  
(Fig. 3) 6 Basic hand postures



(그림 4) 손 제스처 데이터의 다양성  
(Fig. 4) Variations of dynamic hand gestures

이와 같이 다양한 데이터의 특성을 나타내고 있는 동적 손 제스처의 인식을 위해 본 논문에서는 6가지의 손 모양은 P. K. Simpson이 제안한 퍼지 최대·최소 신경망 분류기[8]를 사용하고 7가지 손 운동의 인식은 퍼지 이론을 이용한 특징 해석 방법(feature analysis method)에 의해 이루어진다.

2.1 상태 오토마타에 의한 동작 상태의 인식

연속된 손 제스처를 구분하기 위해 음성인식에 사용되는 은닉 마코프모델을 이용하여 연속된 제스처를 인식하기 위한 시도[10]가 있었으나 이러한 경우에도 사람이 실제로 동작하는 의도적인 제스처와 의미 없는 제스처를 구분하기 어렵다.

본 논문에서는 속력과 속력의 변화량에 의해 '정지 위상', '준비 위상', '스트로크 위상', '동작 진행 위상', '동작 끝 위상'의 5가지 동작 위상(motion phase)을 정의하고 매 샘플 데이터에 대하여 동작위상을 구분한다. <표 1>은 각 위상들을 구분하기 위한 조건들을 나타낸다. 여기서 속력과 속력의 변화는 3차원 x, y, z축에 대하여 다음 식에 의해 구한다.

$$v(k) = \sqrt{x(k)^2 + y(k)^2 + z(k)^2} / \Delta t$$

$$\Delta v = v(k) - v(k-1)$$

, k: sampling step, Δt: sampling time

<표 1> 동작 위상 구분을 위한 속력과 속력변화 조건  
<Table 1> Conditions for segmentation of motion phases

동작위상	동작상태	속력	속력 변화
정지(0)		0	±, 0
준비(1)		+	+, -
스트로크(2)		+++	++
동작(3)		++	+, 0, -
동작끝(4)		+	+, 0, -

-: 속력감소, 0: 움직임 없음, +: 느림, ++: 빠름, +++: 아주 빠름

연속된 제스처에서 개별제스처를 구분하기 위하여 제스처의 상태를 5가지로 정의하였다. 제스처가 수행됨에 따라 변화하는 상태를 표현하기 위하여 이전상

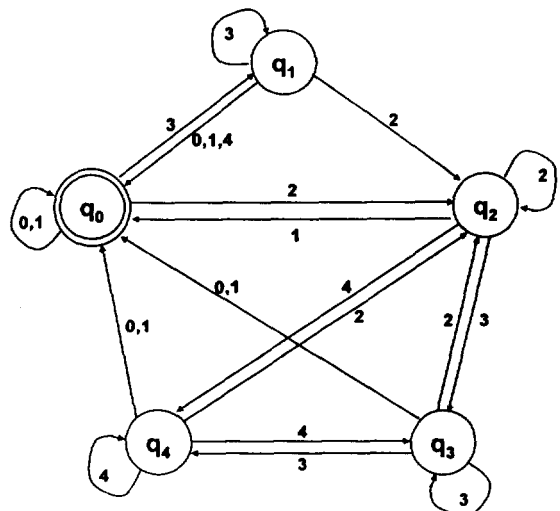
태와 동작 위상에 따라 동작 상태를 변화시키는 상태 오토마타(state automata)[11]를 구현하였다. 상태 오토마타의 구성은 다음과 같다.

$$(E, X, \Gamma, f, x_0)$$

$$E = \{0, 1, 2, 3, 4\}$$

$$X = \{q_0, q_1, q_2, q_3, q_4\}$$

여기서, E는 상태변화를 일으키는 사건 집합으로 속력과 속력변화를 이용하여 구분된 동작 위상이 이에 해당된다. 정지위상에서는 0, 준비위상에서는 1과 같은 심벌의 사건을 할당하였다. 연속된 제스처에서 매 샘플링마다 구분된 동작 위상이 상태 오토마타의 입력 스트림을 이룬다. X는 상태 공간으로 각 동적 제스처를 위하여 정의된 5가지의 동작 상태(q<sub>0</sub>-q<sub>4</sub>)이다. 구분된 상태는 각 제스처에서 다른 의미로 해석되는 동작 상태를 표현한다. Γ는 각 동작상태에서 발생 가능한 사건(동작 위상)이다. 가령, q<sub>2</sub>는 스트로크가 발생한 상태로, 이 상태에서는 정지상태는 발생하지 않기 때문에 Γ(q<sub>1</sub>)에 q<sub>0</sub>(정지상태)가 포함되어 있지 않다. f는 상태 전이함수로 (그림 5)와 같은 상태전이 다이어그램으로 표현할 수 있다. x<sub>0</sub>는 초기상태로 동적 제스처를 위한 초기상태는 q<sub>0</sub>의 정지상태이다. 각 상태는 <표 2>와 같다.



(그림 5) 동적 제스처를 위한 상태 전이 다이어그램  
(Fig. 5) State transition diagram for dynamic gesture

〈표 2〉 동작 상태 설명  
 (Table 2) Description of Current Motion State

동작상태	상태 설명	현재 위상
$q_0$	손 운동이 없거나 의도하지 않은 작은 움직임 상태 또는 손 운동이 끝난 상태	정지 위상(0), 준비 위상(1), 동작끝 위상(4)
$q_1$	동적 순운동이 진행 중인 상태 또는 반복 동작상태	동작진행 위상(3), 동작끝 위상(4)
$q_2$	제스처가 수행될 때의 강한 움직임 상태	스트로크 위상(2)
$q_3$	스트로크가 있었던 움직임 중에 발생한 동작끝 위상, 의미 있는 동작 끝	동작끝 위상(4)
$q_4$	의도적인 제스처 수행이 끝난 상태	준비 위상(1), 정지 위상(0)

모든 동작은  $q_0$  동작 상태에서 수행한다.  $q_4$  동작 상태는 동작 인식을 위한 최종상태로 이 경우 수행된 제스처는 의미 있는 제스처로 간주되어 인식하게 된다. 그렇지 않고 도중에 끝날 경우에는 의미 없는 제스처로 분류된다.

2.2 퍼지 이론을 이용한 특징해석법(Feature Analysis Method)에 의한 손 운동의 분류

각 특징들은 다음의 조건에 의해 선정되었다. 즉, 특징치는 입력 데이터에 대해 정해진 시간(실시간) 내에 계산 가능해야하고 인식뿐만 아니라 운동의 해석에도 사용 가능해야 한다. 또한 모든 운동 형태(7가지)를 분리할 수 있을 만큼의 특징들이 존재해야한다. 여기에 사용된 특징들은 가장 최근의 손의 상대 위치, 손 운동의 누적된 운동 거리, 운동 방향 변화이다. 다음에서 이들 특징에 대해서 상술한다.

(1) 동작 끝 위치(relative Ending Position)

동작 시작점을 기준으로한 동작 끝의 상대적 위치를 나타낸다.

$$EP_{x,y,z} = R_{x,y,z}(k) - R_{x,y,z}(0)$$

$R_{x,y,z}(0)$ : 동작시작위치  
 $k$ : time step(1/15sec.)

(2) 누적 운동 거리(Cumulative length)

손 제스처의 운동 시작이후 매 순간의 각 축에 대

한 각각의 이동 거리를 누적한 값을 나타낸다.

$$CL_{x,y,z}(k) = CL_{x,y,z}(k-1) + L_{x,y,z}(k)$$

$$L_{x,y,z}(k) = |R_{x,y,z}(k) - R_{x,y,z}(k-1)|$$

(3) 방향 변화(Change of Direction) 횟수

왕복 운동의 경우( $D_3$ )를 고려하여 x축에 대한 방향 변화의 횟수를 표시한다.

$$CD_x = CD_x + f(x)$$

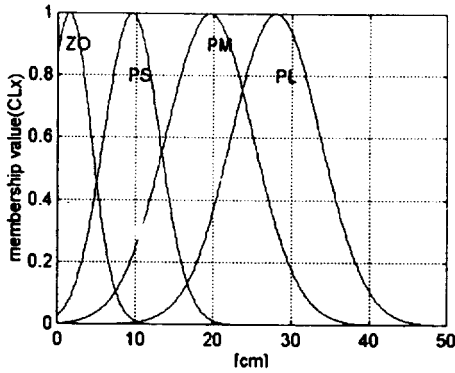
$$f(x) = \begin{cases} 1, & \text{if } [R_x(k) - R_x(k-1)] \times [R_x(k-1) - R_x(k-2)] < 0 \\ 0, & \text{otherwise} \end{cases}$$

이상의 특징들을 이용하여 7개의 기본 방향을 분류하기 위해 퍼지논리를 이용한 분류기를 설계하였다. 즉, 각 손 방향에 대하여 반복 실험을 통해서 특징들을 언어적 변수로 표현하는 규칙을 만들고 이들에 대해 Mamdani의 Max-Min 추론을 하도록 하였다[12].

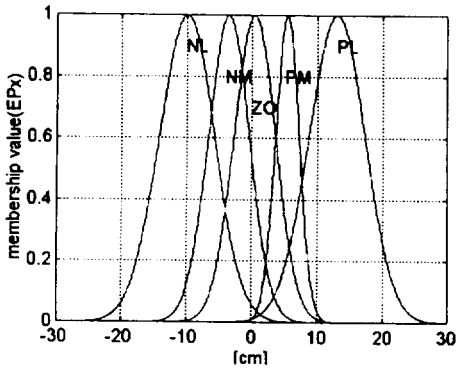
$$A \text{ AND } B = ((x, \mu_{A \text{ AND } B}(x)) | \mu_{A \text{ AND } B}(x) = \min\{\mu_A(x), \mu_B(x), x \in X\})$$

$$A \text{ OR } B = ((x, \mu_{A \text{ OR } B}(x)) | \mu_{A \text{ OR } B}(x) = \max\{\mu_A(x), \mu_B(x), x \in X\})$$

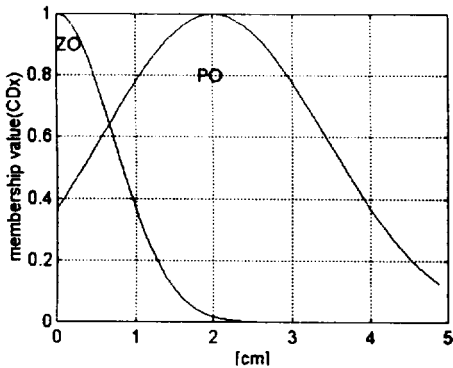
이렇게 하여 구해진 특징 벡터의 영역 구분을 통하여 소속 함수를 구성하였다. (그림 6)은 x축에 대한 소속 함수를 표시한 것으로써 실제 구현에 있어서 계산 속도와 각 동작의 범위 영역을 정하기 용이한 종모양의 소속 함수를 사용하였다.



(a)



(b)



(c)

(그림 6) 소속함수의 구성  
(Fig. 6) Fuzzy Membership Function

퍼지 패턴 클래스는 앞에서 정의된 AND, OR 연산을 이용하여 구현한다. 실제 이 연산을 사용하여 정의된 패턴 클래스의 예를 살펴보면

$D_1$ : (수평 우향운동)의 경우

언어적 기술: 수평 방향의 오른쪽 방향으로의 운동이 있으며 수직 방향이나 전후 방향의 운동이 거의 발생하지 않으며 시작점을 기준으로 수평 우측 방향에서 동작이 끝난다 그리고 동작 중에 방향 변화는 일어나지 않는다.

퍼지 명제 논리식에 의한 표현은

$$(CL_x = PM \text{ or } CL_x = PS) \text{ and } (CL_y = Z0) \text{ and } (EP_x = PM \text{ or } EP_x = PL) \text{ and } CD_x = Z0 \text{ and } CL_z = Z0$$

이고, 다른 손 운동에 대해서도 이와 같은 방법으로 작성할 수 있다.

이와 같이 하여  $D_1$ 에서부터  $D_7$ 까지에 대한 방향 집단의 퍼지 규칙을 구현 할 수 있다. <표 3>은 이와 같

<표 3> Fuzzy Rule Table  
<Table 3> Fuzzy Rule Table

운동 방향	규 칙
$D_1$	$(CL_x = PM \text{ or } CL_x = PS) \text{ and } (CL_y = Z0) \text{ and } (EP_x = PM \text{ or } EP_x = PL) \text{ and } CD_x = Z0 \text{ and } CL_z = Z0$
$D_1$	$(CL_x = PM \text{ or } CL_x = PS) \text{ and } (CL_y = Z0) \text{ and } (EP_x = PM \text{ or } EP_x = PL) \text{ and } CD_x = Z0 \text{ and } CL_z = Z0$
$D_1$	$(CL_x = PM \text{ or } CL_x = PS) \text{ and } (CL_y = Z0) \text{ and } (EP_x = PM \text{ or } EP_x = PL) \text{ and } CD_x = Z0 \text{ and } CL_z = Z0$
$D_2$	$(CL_x = PM \text{ or } CL_x = PS) \text{ and } (CL_y = Z0) \text{ and } (EP_x = NM \text{ or } EP_x = NL) \text{ and } CD_x = Z0 \text{ and } CL_z = Z0$
$D_3$	$(CL_x = PM \text{ or } CL_x = PL) \text{ and } (CL_y = Z0 \text{ or } CL_y = PS) \text{ and } (EP_x = Z0 \text{ or } EP_x = PL \text{ or } EP_x = NL) \text{ and } CD_x = PO$
$D_4$	$CL_x = Z0 \text{ and } (CL_y = PS \text{ or } CL_y = PM) \text{ and } (EP_y = PM \text{ or } EP_y = PL) \text{ and } CL_z = Z0$
$D_5$	$(CL_x = Z0 \text{ and } (CL_y = PS \text{ or } CL_y = PM) \text{ and } (EP_y = NM \text{ or } EP_y = NL) \text{ and } CL_z = Z0$
$D_6$	$(CL_x = PM \text{ or } CL_x = PS) \text{ and } (CL_y = Z0 \text{ or } CL_y = PS) \text{ and } (EP_x = PM \text{ or } EP_x = PL) \text{ and } CL_z = Z0$
$D_7$	$(CL_x = PM \text{ or } CL_x = PS) \text{ and } (CL_y = Z0) \text{ and } CL_y = PS \text{ and } (EP_x = NM \text{ or } EP_x = NL) \text{ and } CL_z = Z0$

이 작성된 클래스를 규칙 테이블로 보여 주고 있다.

2.3 퍼지 최대-최소 신경망 분류기에 의한 손 모양의 분류

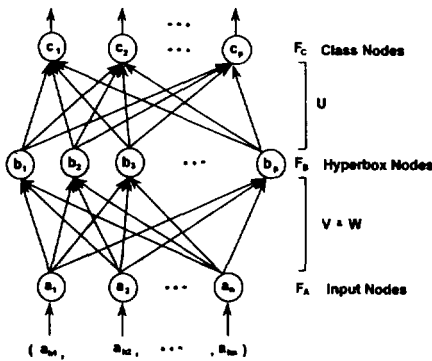
앞에서 설명한 것처럼 사람의 손은 손가락의 굽힘에서 다양한 변화가 존재하고 사람들마다 다양한 크기, 모양의 손가락 구조를 가진다.

본 논문에서는 (그림 3)에 보여진 6가지의 기본 손 모양을 P. K. Simpson의 퍼지 최대-최소 신경망 분류기[8]를 이용하여 실시간으로 분류하고자 한다. 사용된 소속함수는 다음과 같이 표시할 수 있고 (그림 7)은 전체 분류기의 구조를 나타낸 것이다.

$$b_j(F_j, V_j, W_j) = \frac{1}{10} \sum_{i=1}^{10} [1 - f(F_j - w_{ji}, \gamma) - f(v_{ji} - F_j, \gamma)]$$

$$f(x, \gamma) = \begin{cases} 1, & x\gamma > 1 \\ x\gamma, & 0 \leq x\gamma \leq 1 \\ 0, & x\gamma < 0. \end{cases}$$

여기서,  $F_j, j=1, 3, \dots, 9$ 는 손가락의 첫 번째 관절의 굽힘 정도이고  $F_j, j=2, 4, \dots, 10$ 은 두 번째 관절의 굽힘 정도를 나타내고 있다.  $\gamma$ 가 커지면 소속함수는 점점 크리슌(crisp)하게 되므로 본 시스템에서는 첫 번째 관절의  $\gamma(\gamma_1)$ 가 두 번째 관절의  $\gamma(\gamma_2)$  보다 작게 하였다. 손가락의 입력 정보에 의해 이 분류기의 출력은 6가지 집단에 대한 각각의 소속 정도를 나타내고 역치(threshold)를 초과하는 최대값을 가진 집단이 현재 입력된 손 모양의 소속집단이 된다.

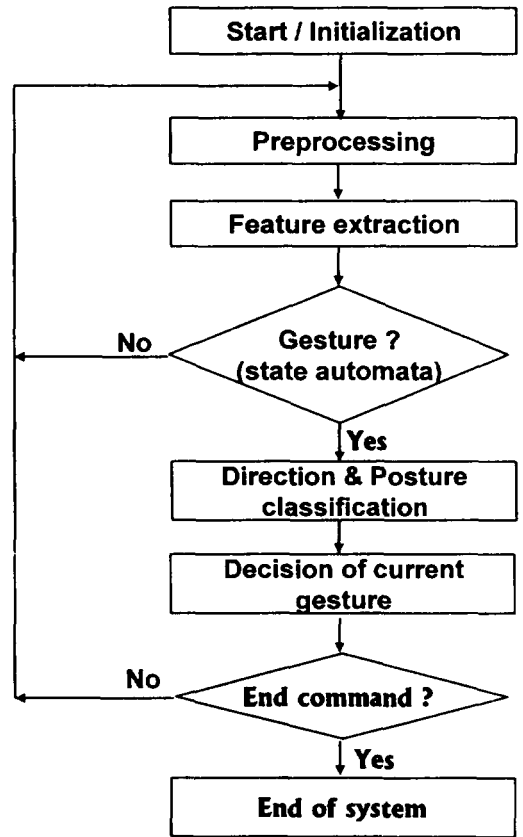


(그림 7) 퍼지 최대-최소 신경망의 구조 [8]  
(Fig. 7) The structure of Fuzzy Min-Max Neural Network [8]

3. 손 제스처 명령의 인식

사용자(signer)가 손 제스처 명령을 시작함에 따라 CyberGlove™로부터 많은 데이터가 시스템에 입력 되게 되고 위에서 설명한 특징 해석법은 이러한 데이터를 특징을 포함하는 작은 데이터 집단으로 만들고 기본적인 7가지 방향을 인식하기 위해 사용되어진다. 모든 동작의 시작은 3차원 공간상의 임의의 위치에서 이루어질 수 있으며 특징 추출과정에서 미소한(1인치 이하) 움직임은 무시되기 때문에 가상 현실공간에서 작업할 때 생기는 사용자의 약간의 손 떨림, 작은 손의 움직임 등은 무시할 수 있게 된다.

미소 범위 이상의 동작이 시작되면 특징 추출과정이 진행되고 이후 앞장에서 설명한 상태 오토마타에



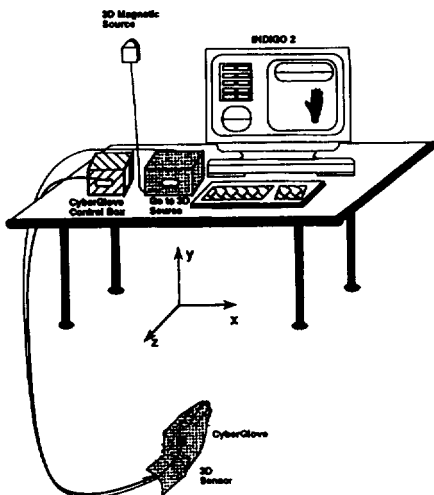
(그림 8) 손 제스처 인식 알고리즘  
(Fig. 8) A flow-chart of overall system



의해 의미있는 동작의 구분이 이루어진다. 그리고 하나의 의미있는 동작이 이루어진 순간에 현재까지 이동한 손의 움직임(방향)에 대한 퍼지 추론이 진행되고 현재의 손의 움직임에 대한 소속함수 값을 얻게 된다. (그림 1)에 있는 '정지', '추적', '잡기' 명령의 경우는 손의 방향에 상관없이 움직임만을 확인할 필요가 있는 명령이므로 7가지의 방향에 대한 정보가 의미가 없다. 일단 현재의 동작에서 방향이 정해지면 이때 손의 모양에 대한 정보가 앞에서 설명한 퍼지 최대-최소 신경망 분류기에 입력되어 6가지 손 모양 중의 하나로 분류가 된다. 이 두개의 결과를 결합함으로써 온라인으로 연속해서 동적 손 제스처 동작을 인식하게 된다. (그림 8)에서 전체 인식 알고리즘의 흐름을 보여준다.

4. 실험결과

(그림 9)는 본 논문의 실험을 위한 가상 현실용 손 제스처 인식 시스템의 구성을 나타낸 것이고 (그림 10)은 실제 모습을 보인다. 전체 시스템은 주(main) 컴퓨터로 INDIGO2 워크스테이션(workstation)을 사용하였고 CyberGlove™는 38,400 baud rates로 RS232 통신을 통해 주 컴퓨터에 13종류의 데이터(10개의 손 굵힘, 3개의 위치)를 보낸다.



(그림 9) 손 제스처 인식 시스템의 구성  
(Fig. 9) The configuration of gesture recognition system

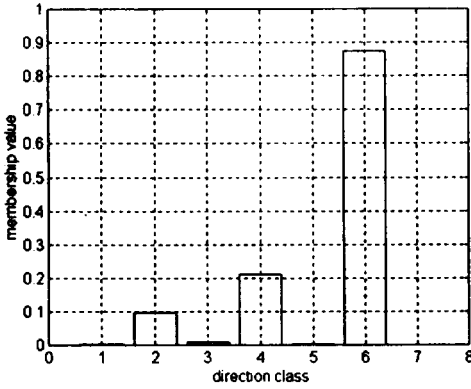


(그림 10) 손 제스처 인식 시스템의 외관  
(Fig. 10) The picture of gesture recognition system

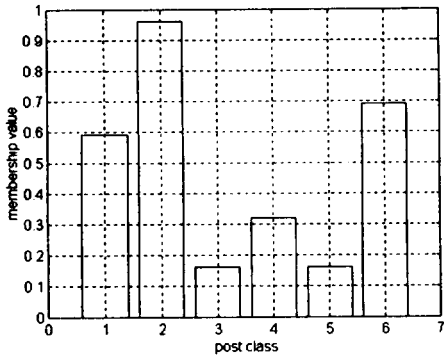
전체 인식시스템은 크게 4가지의 단계로 나누어진 다. 첫째는 CyberGlove™를 착용한 사용자가 동작의 시작 지점을 설정한다. 이때, CyberGlove™로부터의 데이터가 특징 데이터로 변환된다. 두번째는 본 논문에서 설계한 상태 오토마타에 의해 연속적인 동적 제스처에서 의미있는 손 제스처를 분리한다. 세 번째는 7가지의 기본 방향이 퍼지 이론을 이용한 특징 해석법에 의해 분리되고 6가지의 손모양에 대한 분류가 이루어진다. 마지막으로 동적 손 제스처의 결과가 위 2개의 분류결과를 결합함으로써 음성 및 그래픽으로 표현된다.

(그림 11)은 엄지손가락과 집게 손가락을 펴고 손을 앞으로 움직이는 '앞으로(forward)'라는 동작에 대한 인식 결과를 보인 것이다. (그림 11(a))는 손의 방향에 대한 특징 데이터를 이용한 퍼지 추론의 결과를 나타낸 것이다. 3장에서 설명한 퍼지 추론에 방법에 의해 6번째 방향집단이 최대의 소속함수값을 표시하고 방향이 'D<sub>6</sub>'로 정해지고 곧이어 이때의 손가락 정보가 FMMNN에 입력되어 (그림 11(b))에 보여진 결과를 얻는다. (그림 11(b))에서 수평축에 있는 숫자는 (그림 3)에 있는 상용하는 손 모양의 숫자와 관계된다. FMMNN에서 사용된 변수는  $\theta=0.2$ ,  $\gamma_1=4.0$ ,  $\gamma_2=8.0$ 이고  $\delta=0.90$ 이다. (그림 11(b))에서 보는 바와 같이 2번째 집단(P<sub>2</sub>)이 역치( $\delta$ )를 초과하는 최대값을 가지고 따라서 두 결과를 이용하여 이 동작이 '앞으로'임을 정할 수 있다.

5명의 다른 실험자를 대상으로 10가지의 손 제스처



(a)



(b)

(그림 11) 실험 결과 '앞으로'

(Fig. 11) Experimental result on gesture 'Forward'

명령(그림 1)에 대하여 연속적(continuous)으로 10회의 반복 실험을 수행하였고 97.6%의 평균 인식율을 얻었다. <표 4>는 각 제스처 명령에 대한 인식율 결과를 나타낸 것이다. 손 제스처 명령 '오른쪽으로' 경우는 오른쪽방향( $D_1$ )을 뒤쪽방향( $D_7$ )으로 잘못 인식한

경우가 대부분이었으며 사용자가 우측 뒤쪽으로 움직인 경우가 많았다. 또한 '추적' 명령의 경우는  $P_6$  (그림 3 참조)을  $P_1$ 으로 잘못 인식한 경우이다. 이와 같이 인식 오류는 주로 이용자의 손 제스처 동작 속도의 미숙, 동작 중에 발생한 동작 미숙, 그리고 비주기적인 센서의 오동작 등에서도 기인하고 있다.

5. 결론 및 추후과제

본 논문은 컴퓨터와 인간의 상호 접속에서 가장 중요한 매개체가 되고 있는 손 제스처의 인식을 위해 다양하고 복잡한 사람의 손 제스처 중에서 가장 체계적인 한글 수화(KSL)에서 선택한 10가지 손 제스처를 대상으로 연속적인 동작 손제스처에서 상태 오토 마타를 이용하여 의미있는 손 제스처를 구분하고 인식 대상의 확장이 용이한 기본 요소(손의 방향, 손의 모양) 분류기에 의한 인식 알고리즘을 개발하였다. 본 논문에서는 퍼지 이론을 이용한 특징 해석법에 의해 기본 7가지 방향을 실시간으로 분류하고 퍼지 최대. 최소 신경망을 이용하여 기본 6가지 손 모양을 분류하여 온라인으로 손 제스처를 인식하는 시스템을 구현하였다.

본 연구는 서론에서 언급한 일반적인 손 제스처의 인식 시스템(수화 인식, 수신호 등) 구현에도 쉽게 응용 가능하며, 앞으로 인식가능 동작의 확대를 위한 기본 요소의 확장 및 양손에 의한 손 제스처의 확장 시스템을 구현하고자 한다.

참고 문헌

[1] David J. Sturman, *Whole-Hand Input*, Doctoral Dissertation of MIT, Feb. 1992.

<표 4> 손 제스처의 인식율  
<Table 4> Recognition rate of hand gesture

명령	앞으로	뒤로	좌로	우로	정지	위로	아래로	인사	추적	잡기	계(%)
방향분류기(오차 %)	0	2	2	5	0	0	0	6	0	0	1.5
손 모양 분류기(오차 %)	0	0	0	0	0	2	1	0	4	2	0.9
전체 인식 성공률(%)	100	98	98	95	100	98	99	94	96	98	97.6

[2] D. Morris, *Manwatching: A Field Guide to Human Behaviour*, Grafton, 1978.

[3] S. Sidney Fels and G. E. Hinton, "Glove-Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesizer," *IEEE Trans. Neural Networks*, vol. 4, no. 1, pp. 2-8, Jan. 1993.

[4] T. Yamaguchi and et al., "Japanese sign language recognition system using information infrastructure," in *Proc. IEEE Int. Conf. on Fuzzy System, FUZZ-IEEE'95*, pp. 65-66, Mar. 1995.

[5] M. B. Waldron and S. Kim, "Isolated American sign language recognition system for deaf persons," *IEEE Trans. on Rehabilitation Engineering*, vol. 3, pp. 261 - 271, Sept. 1995.

[6] D. Kessler, L. Hodges, et. al, "Evaluation of the CyberGlove™ as a whole hand input device," *ACM Trans. on CHI*, vol. 2, pp. 263-283, Dec. 1995.

[7] Virtual Technologies, Inc., *CyberGlove™ User's Manual*, CA, 1995.

[8] P. Simpson, "Fuzzy Min-Max Neural Networks- Part 1: Classification," *IEEE Trans. Neural Networks*, vol. 3, pp. 776-786, Sep. 1992.

[9] 김 승국, *표준 수화사전*, 오성 출판사, 1993.

[10] T. Starner et al., "Visual Recognition of American Sign Language Using Hidden Markov Models," in *Proceeding of Int. Workshop on Automatic Face and Gesture Recognition*, pp. 189-194, Swiss. Zurich, June 1995.

[11] C. G. Cassandras, *Discrete Event Systems*, IRWIN Boston

[12] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, New York: Plenum Press, 1981.

[13] P. A. Harling and A. Edwards, *Progress in Gestural Interaction*, Proceedings of gesture workshop'96, Springer, 1996.

[14] D. Sturman and D. Zeltzer, "A Survey of Glove-based Input," *IEEE Computer Graphics and Applications*, vol. 14, no. 1, pp. 30-39, Jan. 1994.

[15] J. -S. Kim, W. Jang, and Z. Bien, "A Dynamic Gesture Recognition System for the Korean Sign Language," *IEEE Trans. on SMC*, vol. 26, no. 2, pp. 354-359, April, 1996.

[16] 한국 청각 장애인 복지회 편저, *사랑의 수화 교실*, 수협사, 1995.



**김 종 성**

1989년 2월 경북대학교 공과대학 전자공학과 졸업(학사).

1991년 2월 한국 과학기술원 전기 및 전자공학과 졸업(석사).

1996년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학 박사).

1996년 3월~1997년 1월 한국 과학기술원 전기 및 전자공학과 위촉연구원. (Post-Doc.)

1997년 2월~현재 한국 전자통신 연구원 멀티미디어 연구부 인공지능연구실 선임연구원.

관심분야: 패턴 인식, Machine Vision, Intelligent System(Fuzzy, 인공 신경망), 가상 현실, 재활공학(Rehabilitation Engineering)



**이 찬 수**

1995년 2월 연세대학교 전자공학과 졸업(학사).

1997년 2월 한국 과학기술원 전기 및 전자공학과 졸업(석사).

1997년 3월~현재 시스템 공학 연구소 가상현실 연구실 연구원.

관심분야: 제스처 인식, 휴먼 컴퓨터 인터페이스, 가상현실, 지능 시스템



**송 경 준**

1982년 2월 명지대학교 공과대학 전자공학과 졸업(학사).

1984년 2월 명지대학교 공과대학 전자공학과 졸업(석사).

1982년~1984년 명지대학교 전자공학과 실험조교.

1985년 1월~현재 한국 전자통신연구원 인공지능연구실 선임연구원.

관심분야: 가상현실, HCI, 멀티미디어시스템



**민 병 익**

1982년 한양대학교 졸업(학사)

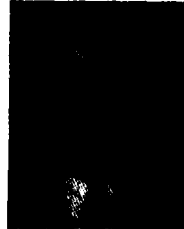
1984년 한국과학기술원 전기 및 전자공학과 졸업(석사)

1992년 한국과학기술원 전기 및 전자공학과 졸업(박사)

1984년~1987년 대림산업 기술연구소

1987년~현재 한국전자통신연구원 인공지능연구실 실장

관심분야: 멀티미디어 시스템, 에이전트 시스템



**박 치 항**

1974년 서울대학교 응용물리학과 졸업(학사).

1980년 한국과학기술원 전산학과 졸업(석사).

1987년 파리 6대학 전산학과(박사).

1974년~1978년 한국과학기술연구소 연구원.

1978년~1985년 한국전자 기술연구소 선임연구원.

1985년~현재 한국전자통신연구원 컴퓨터연구단장.

관심분야: 멀티미디어, 분산시스템, 가상현실, 에이전트, 네트워크 가상 컴퓨팅